The solar virtual observatory from the HESSI perspective

André Csillaghy

for the HESSI Software Team

Space Sciences Laboratory

University of California, Berkeley

HESSI SoftwareTeam

PI: Robert Lin (UCB)

Co-I: Dick Canfield (MSU), Brian Dennis (GSFC), Gordon Hurford (UCB)

Associated personnel: David Ardila (UCB), Markus Aschwanden (LMSAL) Robert Bentley (MSSL), Merrick Berg (GSFC), Martin Fivian (PSI), Chris Johns-Krull (UCB), Säm Krucker (UCB), Tom Metcalf (LMSAL), Jim McTiernan (UCB), Ed Schmahl (UMD), Richard Schwartz (GSFC), David Smith (UCB), Kim Tolbert (GSFC), Greg Walton (UCB), Dominic Zarro (GSFC)

Outline

- HESSI (High Energy Solar Spectroscopic Imager) short presentation
- HESSI involvement in data integration
- Data integration issues and solutions
- Current work

HESSI Science Objective

To explore the basic physics of particle acceleration and explosive energy release in Solar Flares

- Impulsive Energy Release in the Corona
- Acceleration of Electrons, Protons, and Ions
- Plasma Heating to Tens of Millions of degrees
- Energy and Particle Transport and Dissipation

HESSI Primary Observations

- Hard X-ray Images
  - Angular resolution as fine as 2 arcseconds
  - Temporal resolution as fine as 10 ms
  - Energy resolution of <1 keV from 3 keV to 400 keV
- High Resolution X-ray and Gamma-ray Spectra
  - keV energy resolution
  - To energies as high as 20 MeV

HESSI Firsts

- Hard X-Ray Imaging Spectroscopy
- High Resolution Spectroscopy of Solar Gamma-Ray Lines
- Hard X-Ray and Gamma-Ray Imaging above 100 keV
- Imaging of Narrow Gamma-Ray Lines
- High Resolution X-ray and Gamma-Ray Spectra of Cosmic Sources
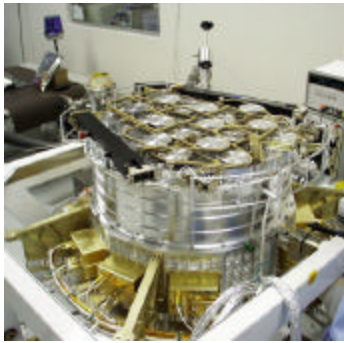- Hard X-Ray Images of the Crab Nebula with 2-arcsecond Resolution

HESSI status

- Vibration and thermal tests passed
- Currently at VAFB for integration into the rocket
- Ship to KSC soon
- Launch March 28, 2001
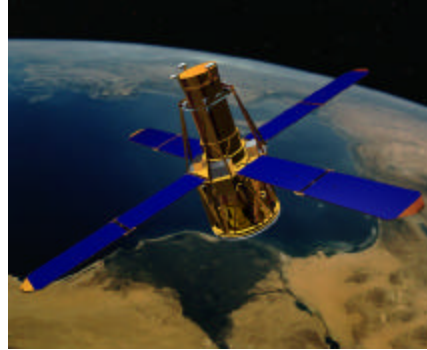- Broadcast on NASA Select
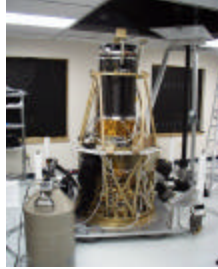
HESSI Imager

HESSI Spectrometer

HESSI in flight

HESSI data management plan

- Open data policy:
  - All data on-line within 24 hours
  - Quicklook data
  - Catalog data
  - Data analysis software on Solarsoft
- User acess:
  - Level-0 telemetry data
  - Calibration database
  - Create their own images / lightcurves / spectra by setting some specific parameters
  - Allow making tradeoffs between energy range, time range, and other characteristic parameters
- 1 GB per day + 30% "value added"

HESSI and data integration

- Most research depends on correlation with other data sets:
  - Radio
  - Soft X-Rays
  - H-α
- Collaboration with 29 observatories
- Need standard access
  - We don't want to make 29 queries for each observation
- Need standard data analysis interfaces
  - We don't want to learn 29 different data analysis software
- Need tools for combination, manipulation of multiple data sets
- Need tools for data mining software
  - similarity searches

Example

- A solar flare occurs at 2001/01/28 10:00:00
- First, check on the web for quick overview
  - Quicklook data
- Then, analysis:
  - HESSI data set
    ```
    o = hsi_image()
    o->plot, time_range = …
    ```
    Will get the data in the (nearest) public archive and plot it
- Then: association with other observations?
  - Ideal:
    ```
    o = map( 'hessi', 'sxt', 'ovsa' )
    o->plot, time_range = [a,b], /LIGHTCURVE
    ```

Issues

Data retrieval: How to *retrieve* data sets from multiple sources transparently ?

Data analysis: How to *integrate* data sets at the analysis level ?

Data retrieval: issues

- General comment: we are not alone
  - Business
  - GIS
  - Issues addressed already
  - Astronomy is behind
- Astronomical data more complex?
  - Calibration
- Technology is available
  - RDBMS
  - ORDBMS
  - Spatial Data Structures (B-Trees, Gridfiles, GiST)
- There is no general schema

Data retrieval: search for a solution

- Not yet another web page
- Need a retrieval system:
  - Location independent: don't transfer the data to another site
  - Scalable: add as many sites as possible
  - Flexible: keep the database management local
- Based on GIS experience
- (O)RDBMS
  - Informix, Sybase, Oracle, ...
- Extensions to:
  - Garlic (IBM/DB2)
  - TSIMMIS (Stanford)
  - Telegraph (Berkeley)
  - Cohera (Payware)

Data federation systems

- Leave individual (local) data archives as they are, and build a management layer on top
- Features (from Cohera, Inc.):
  - Location Transparency
  - Real-time Integration
  - Standard Interface
  - Heterogeneous Data Source Support
  - Flexible Replication
  - Unified Security Model
  - Local Autonomy and Policy Management
  - Intelligent Distributed Query Execution
  - Heterogeneous Semantics Support
- Used in business applications

Data federation systems: current work

- Deploy a data federation system in the context of HESSI
- Involvement:
  - HESSI data archive
  - SXT
  - Big Bear
  - OVSA
  - Max millennium archive
- "Small" virtual solar observatory first
  - Trimmed to output most science as possible
  - Prototype system uses a commercial data federation system
  - Extend later to larger collaborations
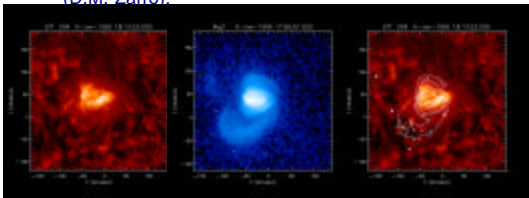
Data analysis: issues

– Integrate the transparent data retrieval into the analysis software
  – e.g. functiion Get(time1, time2) returns observation data beween time t1 and time t2 regardless of the instrument
– Provide common data analysis interface
  – functions Get, GetData, are the same for each instrument
– Provide access to instrument-specific functions
– Provide analysis, and display functions to manipulate data together
  – Zoom, Display

Data analysis: search for a solution

– Object-oriented methods:
  – Each instrument is defined as an object class
  – Each instrument shares a single interface
– More than class extension: design patterns.
  – Standard interfaces
  – Software reuse
  – Clear and consistent design
– Frameworks
  – Implements a basic data analysis unit that can be used over and over again
– Multiple data set analysis:
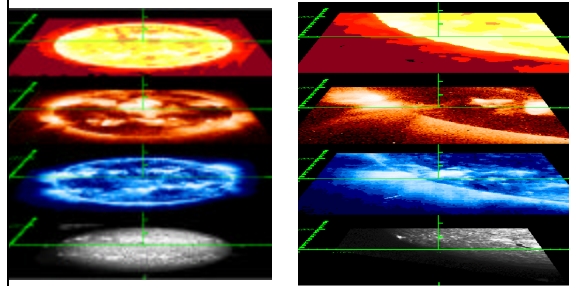 – Display of many data sets simultaneously

Data analysis: current work (1)

– Based on Solarsoft
– Standard frameworks for HESSI
  – All data types accessible through (I.e. implement) a single interface
– Standard mapping system for solar observations (D.M. Zarro):

Data analysis : current work (2)

– Display possibilities for multiple data sets:

Data analysis: current work (3)

– Interface for accessing several data analysis systems
  – Based on the framewoks
– Maps as mentioned earlier, but data analysis functions are active

Conclusion: towards the virtual observatory

– Many technologies are already available
  – Data federation systems
  – Object-oriented design patterns
– Build on these systems rather than start from scratch
– Integrate data retrieval into data analysis
– Technology is analogous for many systems
– Stellar and solar astronomy may collaborate on these issues and reuse software